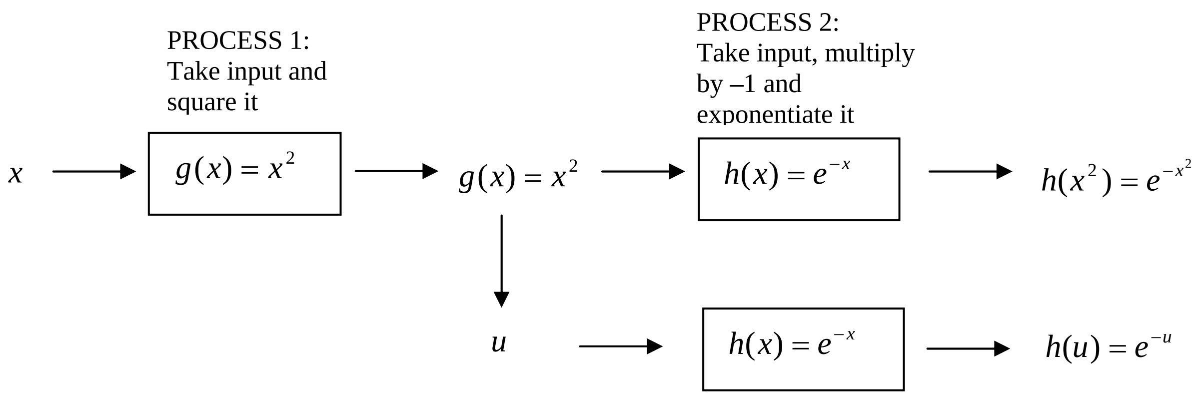
Differentiation rules – The chain rule

The chain rule applies when the function we are considering may be expressed as a function of another function. For example, the function $f(x) = e^{-x^2}$ can be regarded as a function of a function. To see this takes practice, as with everything, but if we let $u = x^2$, then we could legitimately write $f(x) = e^{-u}$

So let us write $g(x) = x^2$ and $h(x) = e^{-x}$, then we can write $f(x) = h(g(x))$. Why is this? Well look at what happens to the input variable x. First of all, it is squared and this is achieved by the "inner" function $g(x)$. The resultant value can then be renamed as u and this becomes the input variable for the "outer" function h. Function h has the instruction "take input, multiply by –1 and the exponentiate it".

If you can see this straightaway, you are probably some kind of mathematical wizard and (more to the point) if you can't, it doesn't surprise me! This really is an acquired skill, but it does come with practise. A useful way of looking at this is via a flow diagram with 2 processes, as shown below:



The downward-pointing arrow in the diagram shows that, after input variable x has been submitted to the first function g, the resultant value can be renamed as a new variable, u, which is in turn submitted to the second function h to yield $h(u)$, which is equivalent to $h(x^2)$ in terms of the original input variable x. Therefore, we can write equivalently all the following:

$$f(x) = h(u) \equiv h(x^2) = h(g(x))$$

So, regarding differentiation from first principles, we take the usual path:

STEP 1: What is $f(x + \delta x)$? Well, in terms of the functions g and h, this is $h(g(x + \delta x))$. Expanding $g(x + \delta x)$, we get

$$f(x+\delta x) = h(g(x) + g'(x).\delta x + O(\delta x^2)) = h(g(x)) + h'(g(x)).\left(g'(x)\cdot \delta x + O(\delta x^2)\right) + O\left(\delta x^2\right)$$

STEP 2: Form the quotient $\frac{\delta y}{\delta x}$.

$$f(x+\delta x) - f(x) = h'\left(g(x)\right)\cdot g'\left(x\right)\cdot \delta x + O\left(\delta x^2\right)$$

and finally,

$$\frac{f(x+\delta x) - f(x)}{\delta x} = h'\left(g(x)\right)\cdot g'\left(x\right) + O\left(\delta x\right)$$

STEP 3: Take the limit as $\delta x \to 0$

$$\lim_{\delta x \to 0} \frac{f(x+\delta x) - f(x)}{\delta x} = \lim_{\delta x \to 0} \left(h'\left(g(x)\right)\cdot g'\left(x\right) + O\left(\delta x\right)\right) = h'\left(g(x)\right)\cdot g'\left(x\right)$$

So, if a function $f(x)$ is defined as $f(x) = h\left(g(x)\right)$, then its derivative is

$$f'\left(x\right) = h'\left(g(x)\right)\cdot g'\left(x\right).$$

Now this form is quite a bit more complicated than (say) the product or quotient rules, because where the meaning of $g'\left(x\right)$ is quite straightforward (it simply means $\frac{dg(x)}{dx}$), the meaning of $h'\left(g(x)\right)$ is not so clear (it certainly does NOT mean $\frac{dh(x)}{dx}$).

Example

The best way to explain this form is by using an example – let's use the one we considered earlier, namely $f(x) = e^{-x^2}$. This is broken down so that the inner function is $g(x) = x^2$ and the outer function is $h(x) = e^{-x}$.

Note that the input variable to the outer function is x^2 and that specifically, it will ***never*** be x, so let's call this variable u. Then let's write out all the relationships that we know to be true:

$$\left.\begin{aligned} u = g\left(x\right) = x^2 \\ f\left(x\right) = h\left(g\left(x\right)\right) = h\left(u\right) = e^{-u} \end{aligned}\right\}$$

Differentiating the parts, we get

$$g'(x) = \frac{dg(x)}{dx} = \frac{du}{dx} = 2x$$

$$h'(g(x)) = \frac{dh(g(x))}{dg(x)} = \frac{dh(u)}{du} = -e^{-u} = -e^{-x^2}$$

which finally gives the derivative of f as:

$$f'(x) = h'(g(x)) \cdot g'(x) = -2xe^{-x^2}$$

The first differentiation is completely straightforward, but the second is not. What we must remember is that, in the second differentiation, we differentiate with respect to the input variable of function h, and not with respect to x. This is why I recommend renaming the input variable of h to u.

As this rule is harder to understand than other ones, it is especially important to practice using this rule to familiarise yourself with the mechanics involved. Let's try another example.

Example

Suppose $f(x) = \sin^2(x)$. What is $f'(x)$?

Well first of all, it is important to realise that $\sin^2(x)$ means $(\sin(x))^2$. So think about what happens to the input variable x. First of all, its sine is taken, providing us with our inner function: $g(x) = \sin(x)$.

Now we make our substitution of variables, so that the output from g, and hence the input to outer function h is the new input variable u. Note that $u = \sin(x)$.

After the sine is taken, the resultant value is squared, giving us the definition of our outer function as $h(u) = u^2$.

So, we have

$$\left.\begin{aligned} u &= g(x) = \sin(x) \\ f(x) &= h(g(x)) = h(u) = u^2 \end{aligned}\right\}$$

Differentiating the various parts, we get

$$g'(x) = \frac{dg(x)}{dx} = \frac{du}{dx} = \cos(x)$$

$$h'(g(x)) = \frac{dh(g(x))}{dg(x)} = \frac{dh(u)}{du} = 2u = 2 \cdot \sin(x)$$

So, the product of these gives the derivative of f as

$$f'(x) = h'(g(x)) \cdot g'(x) = 2 \cdot \sin(x) \cdot \cos(x) = \sin(2x)$$

Example

What is the derivative of $f(x) = \cos(3x^2 - 1)$ at $x = 1$?

The inner function will be the argument of the cosine and the outer function will be the cosine itself, so:

$$\left.\begin{aligned} u = g(x) &= 3x^2 - 1 \\ f(x) = h(g(x)) = h(u) &= \cos(u) \end{aligned}\right\}$$

Differentiate the parts:

$$g'(x) = \frac{dg(x)}{dx} = \frac{du}{dx} = 6x$$

$$h'(g(x)) = \frac{dh(g(x))}{dg(x)} = \frac{dh(u)}{du} = -\sin(u) = -\sin(3x^2 - 1)$$

giving finally:

$$f'(x) = h'(g(x)) \cdot g'(x) = -6x \cdot \sin(3x^2 - 1).$$

Now, the question asked for the numerical value of the derivative at $x = 1$, so we must substitute this value into the expression for the derivative, thus:

$$f'(1) = -6 \times 1 \cdot \sin(3 \times 1^2 - 1) = -6\sin(2) = -5.456$$